Press Release - 3 December 2025

Lack of safeguards is a major gap in the National AI Plan

"The plan takes a 'whack-a-mole' approach to managing AI risk, leaving Australians exposed. Its reliance on 'encouraging' responsible practices from Big Tech companies is out of step with leading jurisdictions overseas and ignores the evidence that action is needed now," warns **Devon Whittle, Global Shield's Australia Director.**

Canberra – Without mandating safeguards, the Australian government's National AI Plan leaves major gaps in Australia's approach to regulating the safe development, deployment, and use of artificial intelligence.

Instead of moving forward with mandatory obligations, such as requiring large AI developers to be transparent with the government and the public, Australians are left to rely on Big Tech companies adhering to voluntary guardrails and individual regulators playing catch-up as harms materialise.

"We cannot build a safe and secure AI ecosystem on 'encouragement'," said Devon Whittle, Global Shield's Australia Director. "Rather than just 'promoting' the use of systems that are transparent and fair, we need minimum and mandatory standards along with transparency requirements to ensure that is the case."

Mandatory safety standards and reporting requirements are common in other high-risk industries. Manufacturers once offered seatbelts as optional equipment in cars, until they were made mandatory by law.

The Plan does usefully recognise that AI will intensify and create new national security threats, including from biological risk and to Australia's critical infrastructure. It also rightly commits the government to integrating consideration of major AI incidents into Australia's crisis management arrangement.

"It is encouraging to see AI being treated as a national security and crisis management issue, not just an innovation and productivity boon," Mr Whittle said. "The new AI Safety Institute could also make a real difference here with focused work on biosecurity and cybersecurity being a natural fit

given Australia's existing expertise and leadership in these areas. The priority should be to pick one or two of these and do them extremely well."

These advances are welcomed, but are overshadowed by the plan's otherwise largely reactive approach to risk and harm. The government states it will step in to regulate only after specific threats or harms become apparent. This 'wait and see' posture is out of step with what is already known about advanced AI models and their potential to enable large-scale harm, including through large-scale cyberattacks and accelerated misinformation and disinformation campaigns.

"We welcome the government's commitment to act when harms emerge or risk requires action. We already have enough evidence to justify targeted guardrails now – especially for the largest AI developers and the most consequential uses," said Whittle. "We need to avoid playing regulatory 'whack-a-mole', with regulators constantly scrambling to react without a foundational framework to guide them."

By reversing course on mandatory guardrails for AI, Australia is also diverging from international leaders. Jurisdictions including the EU, California, New York, Vietnam, and Malaysia are all moving ahead with concrete regulations. At the same time, polling shows Australians have low trust in AI and support stronger action from the government.

"Australians deserve the same protections as their counterparts overseas," said Whittle. "The government must provide this through targeted legislation – including mandatory reporting and safety obligations on the largest AI developers – so we have basic protections in place while capturing the economic benefits of AI."

ENDS

About Global Shield

Global Shield is an independent non-profit advocacy organisation dedicated to reducing global catastrophic risk. Our mission is to help governments and leaders enact and implement effective policies to reduce the risk of global catastrophe from all hazards. www.globalshieldpolicy.org

Contact us: australia@globalshieldpolicy.org / marvin.meintjies@globalshieldpolicy.org